DOCUMENT RESUME

ED 336 493 UD 028 266

AUTHOR Karweit, Nancy L.
TITLE Repeating a Grade: Time To Grow or Denial of Opportunity? Report No. 16.
INSTITUTION Center for Research on Effective Schooling for Disadvantaged Students, Baltimore, MD.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE May 91
CONTRACT RI17R90002
NOTE 20p.
PUB TYPE Information Analyses (070)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Failure; Comparative Analysis; Educational Research; Elementary Education; *Grade Repetition; *High Risk Students; Literature Reviews; *Low Achievement; Meta Analysis; *Research Problems; Research Reports; Student Promotion
IDENTIFIERS *Social Promotion

ABSTRACT

This report examines the limitations of research that has been conducted on the effects of retention and social promotion. It reexamines seven research studies that provide strong evidence based on strong methodology that includes the following three characteristics: (1) the studies identify the basis of comparison, that is, whether they are comparing the achievement of retained versus promoted students at the same age or same grade level; (2) the studies identify the specific kind of educational program that students receive after either retention or promotion; and (3) the studies examine the long-term effects of retention and promotion. It is concluded that neither retention nor social promotion is a satisfactory response to the need to provide appropriate instruction for low performing students. Two tables present summaries of the studies and their effect sizes. A 29-item list of references is included. (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

SCOPE OF INTEREST NOTICE

The ERIC Facility has assigned this document for processing to: UD PS

In our judgment, this document is also of interest to the Clearinghouses noted to the right. Indexing should reflect their special points of view.

ED 336493

CDS

THE JOHNS HOPKINS UNIVERSITY

REPEATING A GRADE
Time To Grow
Or Denial of Opportunity?

Nancy L. Karweit

Report No. 16

May 1991

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Blanket

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

UD 028 266

CENTER FOR RESEARCH ON EFFECTIVE SCHOOLING FOR DISADVANTAGED STUDENTS

2 BEST COPY AVAILABLE

National Advisory Panel

Beatriz Arias, College of Education, Arizona State University

Mary Frances Berry, Department of History, University of Pennsylvania

Anthony S. Bryk, Department of Education, University of Chicago

Michael Charleston, Educational Policy Studies, Pennsylvania State University

Constance E. Clayton, Superintendent, Philadelphia Public Schools

Edmund Gordon (Chair), Department of Psychology, Yale University

Ronald D. Henderson, National Education Association

Vinetta Jones, Dean, Education and Urban Affairs, Morgan State University

Arturo Madrid, Director, Tomas Rivera Center

Hernan LaFontaine, Superintendent, Hartford Public Schools

Peter Stanley, Director, Education and Culture Program, Ford Foundation

William Julius Wilson, Department of Sociology, University of Chicago

Center Liaison

Harold Himmelfarb, Office of Educational Research and Improvement

REPEATING A GRADE

Time to Grow or Denial of Opportunity?

Nancy L. Karweit

Grant No. R117 R90002

Report No. 16

May 1991

Published by the Center for Research on Effective Schooling for Disadvantaged Students, supported as a national research and development center by funds from the Office of Educational Research and Improvement, U.S. Department of Education. The opinions expressed in this publication do not necessarily reflect the position or policy of the OERI, and no official endorsement should be inferred.

Center for Research on Effective Schooling for Disadvantaged Students
The Johns Hopkins University
3505 North Charles Street
Baltimore, Maryland 21218

The Center

The mission of the Center for Research on Effective Schooling for Disadvantaged Students (CDS) is to significantly improve the education of disadvantaged students at each level of schooling through new knowledge and practices produced by thorough scientific study and evaluation. The Center conducts its research in four program areas: The Early and Elementary Education Program, The Middle Grades and High Schools Program, the Language Minority Program, and the School, Family, and Community Connections Program.

The Early and Elementary Education Program

This program is working to develop, evaluate, and disseminate instructional programs capable of bringing disadvantaged students to high levels of achievement, particularly in the fundamental areas of reading, writing, and mathematics. The goal is to expand the range of effective alternatives which schools may use under Chapter 1 and other compensatory education funding and to study issues of direct relevance to federal, state, and local policy on education of disadvantaged students.

The Middle Grades and High Schools Program

This program is conducting research syntheses, survey analyses, and filed studies in middle and high schools. The three types of projects move from basic research to useful practice. Syntheses compile and analyze existing knowledge about effective education of disadvantaged students. Survey analyses identify and describe current programs, practices, and trends in middle and high schools, and allow studies of their effects. Field studies are conducted in collaboration with school staffs to develop and evaluated effective programs and practices.

The Language Minority Program

This program represents a collaborative effort. The University of California at Santa Barbara is focusing on the education of Mexican-American students in California and Texas; studies of dropout among children of immigrants are being conducted at Johns Hopkins, and evaluations of learning strategies in schools serving Navajo, Cherokee, and Lumdee Indians are being conducted by the University of Northern Arizona. The goal of the program is to identify, develop, and evaluate effective programs for disadvantaged Hispanic, American Indian, Southeast Asian, and other language minority children.

The School, Family, and Community Connections Program

This program is focusing on the key connections between schools and families and between schools and communities to build better educational programs for disadvantaged children and youth. Initial work is seeking to provide a research base concerning the most effective ways for schools to interact with and assist parents of disadvantaged students and interact with the community to produce effective community involvement.

Abstract

This report examines the limitations and inadequacies of research that has been conducted on the effects of retention and social promotion, and re-examines the research studies that provide strong evidence based on strong methodology that includes three characteristics: (1) the studies identify the basis of comparison, that is, whether they are comparing the achievement of retained vs. promoted students at the same age or at the same grade level; (2) the studies identify the specific kind of educational program that students receive after either retention or promotion; and (3) the studies examine the long-term effects of retention and promotion. The report concludes that neither retention nor social promotion are satisfactory responses to the need to provide appropriate instruction for low performing students.

Introduction

Perhaps no decision about a child's education is fraught with as much emotion and confusion as the decision to promote a low achieving child to the next grade or to retain that child in the same grade. Current research is vividly pointing out the negative consequences of grade retention (Shepard and Smith, 1989). Despite the research results pointing to negative effects and the costs to districts of grade retention, the practice continues. In fact, retention in some areas (specifically kindergarten) appears to have dramatically increased during the 1980s. In many urban districts, the cumulative effect of grade retention is such that 1 in 2 students will repeat a grade by the third grade.

Historically, grade retention has been associated with later lack of success in school and with eventual dropping out of school (Grissom and Shepard, 1989). Disadvantaged students, males and minorities are more likely to be retained in grade. Grade repetition is a major, commonly occurring event in the educational histories of less successful students.

This report provides a brief overview of issues involved in retention/social promotion policies, describes the limitations and inadequacies of research on the issues, and re-examines the research studies which provide the best evidence about the effects of grade retention on students. A final section discusses how the research on grade retention is linked to general efforts at school reform and attempts to cope with student diversity in the classroom.

Overview of Issues in Grade Retention and Social Promotion

The need to make decisions about promotion and retention arose with the introduction of age-graded schooling in America in the 1840s. With the introduction of graded classes, the question of standards for promotion from grade to grade became an issue.

Throughout the years, toughening or loosening the standards has reflected the political and reform climate of the particular era. In times when schools are under pressure to improve performance - for whatever reason - there is an inevitable emphasis on standards and on the need to tighten requirements for promotion from grade to grade. When schools lack credibility in the eyes of the public, tightening standards is one very public way by which schools can appear to be responsive.

Shepard and Smith (1989) find that the emphasis on social promotion or on retention has varied markedly across the history of public schooling in the United States. In the 1800s rates of grade repetition were extremely high, affecting as many as 70 percent of all students in any one year. There is historically, as well as currently, little systematic evidence on rates of retention, only illustrative examples. For instance, in Iowa in the early 1900s, about 50 percent of the students were retained each year. This figure declined to about 25 percent in the 1930s and eventually to around 10 percent in the 1960s. Also, retention rates varied widely across the United States at any given time. A 1909 study by Ayres indicated that one Massachusetts school district retained 7.5% while a Tennessee school district retained 75.8% of its students each year.

In the 1930s educators recognized that grade repetition might endanger students' social and emotional development, which gave rise to the practice of social promotion. As a result of this policy, students were passed on to the next grade even if they were not ready for the work. Thus the proportion of overage students at each grade level declined from 1918 until 1952. However, providing homogeneous groups for instruction -- which was accomplished by retention -- was simply met by other avenues. Grouping and tracking practices rose accordingly, and dropout rates increased.

The current educational reform movement has seen an increased focus on standards and corresponding increased rates of grade repetition. The *Nation at Risk* (1983) report, which captured the spirit of the reform movement, specifically called for increasing attention to standards and advancement to the next grade on the basis of academic progress and not on age. Nineteen states have since established specific standards for grade promotion and graduation requirements. That such standards lead to greater rates of retention is exemplified by the experience of the

Atlanta public schools. In 1981, after instituting minimum competency requirements, the retention rate was four times the rate in 1980.

The current *cumulative* rates of grade repetition are in many instances as high as they were before social promotion became popular (Shepard and Smith, 1989, Table 3). Many states now have annual rates of grade repetition of about seven percent. If most students are not retained in more than one grade, this rate of seven percent in one particular year generates cumulative rates of at least 50 percent, which is similar to the rates prior to the advent of social promotion.

The Basis for Retention as a Policy

Two themes serve as the basis for retention as educational policy. The first theme is most often encountered in decisions regarding retention or extra-year placements in kindergarten and first grade. The basis for retention at this grade level is typically student immaturity. Students given the "gift of time" - repeating the same grade - have an opportunity to mature. Such authors as Gesell (1982) and Ames (1966, 1980) advocate testing to determine a child's developmental level and placement of the child in school on the basis of his or her developmental, not chronological, age.

This viewpoint has been particularly influential in pre-kindergarten and kindergarten years, where students are retained because of "immaturity" or simply not being behaviorally ready for school. This philosophical approach specifically denies that intervention to improve student maturation or other deficits is possible or desirable. These programs follow a maturational philosophy in which readiness for schoolwork is a quality which unfolds on its own timetable, not to be rushed or pushed.

The second theme is that low achievement is caused by a lack of exposure to the material being taught and thus can be remediated by recycling the student through the material. In this theme, student "failure" is not deemed a failure of the educational system, but of the student. The student needs more time to mature or additional time to go through the material. Rarely is it assumed that the approach or content is inappropriate for the learner; rather, it is assumed that the learner is inappropriate for the material being presented.

A policy of recycling a student through the material also assumes that learning is a linear process and that mastery of content at one level depends upon mastery at a previous level. It assumes that what is to be learned must be broken down into sequential steps and that mastery of the whole is only achievable by mastery of the parts.

Failure is also often seen as a positive experience for children because it motivates them to succeed. Fear of failure is often offered as a reason to have grade repetition as a policy.

Who is Retained?

The numerous studies of grade retention provide a consistent profile -- the retained student is more likely to be male, to be younger than his classmates, to come from a lower socio-economic family background, to be black or Hispanic, to be a behavior problem, and to be immature. Also, students are more likely to be retained in the South than in any other region. Students are more likely to be retained at specific transitional points, such as kindergarten or first grade (school entry) or grade six (exit from elementary and entry into middle school), or grade nine (high school entrance).

Research and Its Limitations

Although literally hundreds of reports, reviews, dissertations and essays have examined the benefits and drawbacks of grade repetition, they provide limited evidence about whether grade repetition will actually help or harm a student. Four issues limit the utility of most of the research. The first limitation is the design of the research (Jackson, 1975). The second limitation is that studies fail to identify the *basis of comparison* or they improperly combine and aggregate results that use different bases of comparison.

The third limitation is that the studies fail to identify the *educational practice(s)* called "retention" or they inappropriately combine studies which use very different practices. For example, meta-analyses often combine the effects of studies of programs which simply recycle students with studies of programs which provide specialized remedial assistance.

The fourth limitation is that many studies fail to examine the *longitudinal effects* of retention, which could help determine why early grade retention is such a powerful indicator of later school

failure. Also, a good deal of research focusing on educational practices in the early years has documented the fade out phenomenon -- that is, initial effectiveness followed by gradual loss of effects. Retention seems to exhibit this phenomenon, but this needs to be clarified through longitudinal studies.

The design limitations discussed by Jackson (1975) are typically acknowledged in discussions of grade retention, but the difficulties imposed by failure to identify the basis of comparison, to describe the educational practice under study, and to look for longitudinal effects are not readily discussed in the current literature.

Research Design

Jackson's (1975) review of grade repetition categorized studies by their methodology and pointed out the influence of design on the results of the studies. Jackson reviewed 44 studies conducted from 1911 through 1973. He classified the available studies into three study designs:

Design Type I: Studies which compared the retained students with promoted students. This type of design is biased in favor of promoted children. Those students who are promoted are not likely to have the same academic and social problems as those who are being retained or else they would not be promoted. The existence of pre-existing differences in the two populations prior to the event of retention which are not controlled for in the analyses invalidates the results of Type I design.

Design Type II: Studies which compared retained students before and after retention. This type of design is biased in favor of the retained students. This design does not compare retained and promoted students, but shows only the effects of spending two years learning material which was intended to be learned in one year. It is reasonable to expect that students will make more progress the second time around, so this type of study is biased in favor of finding effects for retention. These studies do not control for other factors which positively affect growth.

Design Type III: Studies which randomly assigned equivalent students to promotion or retention. These studies provide the best evidence for effectiveness. Jackson found only three in this category (Cook, 1941; Farley, 1936; Klene and Branson, 1929) and the most recent was published in 1941. Cook (1941) looked at the results of retention in a year-long study of students in grades 1-7 and found no significant differences. Farley (1936) examined the effects of retention in a semester-long study and found that the promoted students outperformed the retained students. Finally, Klene and Branson (1929), in their study of retention and promotion of second through sixth graders, found that the promoted students fared better than the retained students, but did not report levels of significance or enough data to allow computation of levels of significance.

Thus, these three key studies produced only one significant effect and that favored the promoted group over the retained. Jackson interpreted these results to mean that no valid research results showed the positive effect of retention. Much the same conclusion has been reached by more recent research syntheses (Holmes, 1986; Holmes and Matthews, 1984) which have included studies published since Jackson's review. Even more recent studies, however (such as Shepard and Smith, 1989), not only conclude that grade repetition has not shown a benefit, but that it is harmful.

Basis of Comparison

The current meta-analyses of grade repetition have combined studies that focus on different bases of comparison. Students who are retained in grade, by virtue of their retention, will spend more time to attain the same grade in school as their same agemates. The question is whether the progress of students should be compared after they have spent the same time in school (same time, but different grades) or after they are in the same grade (same grade, but different time). Although recent meta-analyses (Holmes, 1989; Shepard and Smith, 1989) present results separately for same age and same grade comparisons, they ultimately combine the results across comparisons and treat the differences in effects as a methodological, not substantive, issue.

Educational Practice

The approximately 800 studies of grade repetition essentially examine the same questions again and again, albeit in different decades, different districts, and for different grades: Does grade repetition hurt or harm students? Are at-risk students better off with social promotion or retention? But embedded in these studies are very different educational practices that are all being called "retention." Transition room placement, developmental kindergartens, partial grade retention, complete grade repetition, and alternative pro-

grams are just some of the treatments reviewed collectively as grade repetition.

At least four types of educational practice have been grouped together under the heading grade repetition:

- *recycling*: repeating the grade, but receiving no additional resources or special program

- *alternative after failure:* repeating the grade, but receiving additional help and special programs

- *alternative pre failure:* being placed in an additional-year program prior to actual failure

- *partial promotion:* being failed or promoted in certain subjects only

Most studies of the effect of grade repetition have mixed the types of programs under one generic name and not looked for distinct effects by type of retention employed. Often it is not possible to tell what happened in the year of retention. Given the lack of clarity of the "treatment," it is suspect to make conclusions about effectiveness or ineffectiveness. If no study ever found a positive effect, we might conclude that all things lumped under the heading "retention" were ineffective. But some studies do find positive effects, so we need to examine the variation in findings to see if there are practices in common across these positive studies.

Longitudinal Comparisons

Studies also need to consider long-term effects of retention. Many educational practices, such as participation in certain preschool programs, have been shown to have an effect in the first year following the treatment that eventually levels out or washes out completely. The long-term effects of retention need to be looked for as well as more immediate effects.

Best Evidence on the Effects of Retention

The intention of this paper is not to review again every study done on grade repetition to arrive at a new estimate for the effect for grade retention. Instead, the intention is to reexamine those studies which provide the best evidence of the effects of retention keeping in mind the limitations discussed by Jackson (1975) and the additional difficulties just discussed -- the need to identify the basis of comparison and the educational practice being examined and the need to look for short term and long term effects.

The consensus of several extensive reviews of grade retention (Jackson, 1975; Holmes, 1986,1989; Holmes and Matthews, 1984) is that there is not a positive effect for grade retention on academic achievement or on student personal adjustment. Holmes (1989) summarizes the position aptly in a recent meta-analysis: "The weight of empirical evidence argues against grade retention." As Holmes and Matthews concluded in 1984, "Those who continue to retain pupils at grade level do so despite cumulative research evidence showing that the potential for negative effects consistently outweighs positive outcomes."

Shepard and Smith take a more strident view, asserting that "retentions do nothing to promote the achievement of the affected individuals or the average of the group as a whole and because the disadvantaged and minority children are most apt to be affected, retention should best be thought of as educational waste and a denial of life chances to those who most need the benefits of education. Retention has high cost and virtually no value, save the public relations advantages for the schools." (p.235).

Despite the conclusions of these and other researchers, districts continue to retain children in grade. Although retention in grade does benefit districts by raising average grade level test scores (Walker and Levine, 1988), districts probably continue to retain because teachers and principals and parents believe it is in the best interest of the individual child. Shepard and Smith (1989) suggest that districts continue to retain because they are unable to carry out experiments to compare a particular child's progress when retained with what it would be had he been promoted. That is, school districts do not have a valid basis for comparison.

What is the Basis for Comparison?

School districts are not the only parties potentially confused about the basis of comparison. Most research poorly identifies who and on what basis the comparison is being made. Are they comparable children? Are the comparisons made after equal time in school and unequal grade? Or, are the comparisons made after unequal time in school and equal grade? Are comparisons made at one time point or multiple ones? Are affective as well as achievement measures compared?

The meta-analysis conducted by Holmes (1989) provides some useful data to address the comparison basis issue. He presents average effect sizes for studies which compared the retainees and non-retainees after equal time in school (same age, different grade) and after equal grade in school (same grade, different ages). The first comparison indicates what happened to the achievement after retainees have spent another year in school, compared to their same age peers who are now in a different grade. The second comparison indicates how the retained children performed when they reached the same grade, although they are now a year older.

Holmes determines that the first comparison (same year) favors the promoted group. The average effect size is -.45, indicating that the retained group is nearly a half a standard deviation below their same age peers (who are also a grade ahead of them). This finding is consistent with Jackson's predictions about the manner in which methodology affects outcomes. Although the control groups were matched on a number of factors, the two groups no doubt continue to be different on a number of factors affecting the decisions to be promoted or retained. No matter how many factors students are matched on, there are always unmeasured factors at work which may favor the promoted group. The negative findings for retention in the same year comparison include some unknown amount that is due to unmeasured initial differences.

When one looks at the same *grade* comparison, a different pattern is presented. The retained students are higher in the first year (average effect size = +.25), but the effects diminish over time until by the third year, they are no longer discernible. There is an initial boost followed by a fade away phenomenon.

The year and grade comparisons are both interpreted by Holmes to indicate that retention is not an effective policy. The same year comparisons clearly are negative and the same age comparisons, while positive, fade away in time. Holmes combines these effect sizes into a global estimate of the effect of retention.

But, does it make sense to combine studies which measure same grade and same year comparisons? What educational significance should be afforded the same grade vs. same year comparisons? In part, the answer to that question depends upon the goal one sets for retention. Are students expected to be remediated by the retained year and be up to the same level as their agemates? Or, are students expected to need additional time and at some point to be at the same level as their classmates? This question of intention is of course central in the evaluation of retention as a policy. It is also possible that there are different answers in different districts and at different levels of schooling. We look now at selected studies which can be used to address some of these issues.

Six studies among the twelve identified by Holmes (1989) provide the base of research evidence to be considered here. These studies allow identification of the basis of comparison as well as the treatment and focus on longitudinal effects.

INSERT TABLE 1

Table 1 provides the effect sizes for the various grade/subject/comparison basis combinations.

Peterson, DeGracie and Ayabe (1987) report on a longitudinal study of the effect of grade retention in the Mesa Public Schools. The achievement effects of retention were compared for students matched on sex, age and previous CAT score. Achievement for four years in the areas of reading, language and math were compared. The program combined retention with remedial assistance.

They carried out comparisons with same year and same grade peers separately for students by grade. The first comparisons look at students in the same year, but who are in a different grade. The basis of comparison is identified in the table as *year*. Because this was a longitudinal study, the students are being compared at distinct time points as well, here identified as T1, T2, T3, T4 and T5. The study followed the students for three years, so effect sizes for differences between the retained and promoted groups are presented for T1, T2 and T3. This says that the effect size favoring retention using same year comparison for grade 1 students on the CAT reading was 1.06. The average effect size for the nine comparisons for grade 1 across the different subtests on the CAT was +.42. When the students were compared on these same subtests at the same grade in school (see the fourth block of results), the average effect size across the three subtests was +.70.

The *same grade* grade comparisons result in larger effect sizes than the *same year* comparisons. But both are positive results, indicating a positive effect for retention. There is the typical

diminuition of effects in time. The program was a retention + individual plan program, which suggests that the positive effect in the year comparisons could be the result of more appropriate instruction. Whether looking at grade or age comparisons, the effects of retention were substantial and *positive*. However, both low-performing promoted and retained students were still performing below district level.

The importance of longitudinal studies in examining the effects of grade retention is underscored by Baenen (1988), who followed 243 matched retained or promoted students in a five year study (see second entry Table 1). The basis of this comparison was when students were in the same grade. For the first three years, the results favor the retained group over the promoted group. The effect size across reading and math for the same grade comparisons for the first three years is .35. It is important that they extended the study past this time, however, as these positive effects eventually fade and in fact become negative. The average effect size across all five years is .15. The students received additional remedial services in this program.

Dobbs and Neville (1967) matched 30 students who were retained in the first grade on sex, race, SES and reading achievement with students who were promoted. They compared the achievement in reading and math on the Metropolitan Readiness test at the same year and once the students were in the same grade. The average effect size for the same year comparison was -.75, indicating that the promoted group, which was in the next grade, had higher achievement scores. The same grade comparisons, carried out over two time points, show an average effect size of 1.36 favoring the retained students. However, this effect size is inflated by the small number of cases and by the short duration of the study. The educational practice employed was reexposure to the same material.

Coffield (1954) compared the achievement of current seventh graders who had failed at some point during grades 3-7 with the achievement of students matched on the ITBS in the year prior to the retention. He compared the achievement of the two groups in the year of the retention (same year comparison) and when the students were in the same grade (same grade comparison). In the report, analyses were conducted by the year in which failure occurred. There was not a significant difference across year failed so we averaged the effect sizes to come up with a -.77 overall effect size for the same year comparison and a -.17 effect size for the same grade comparison. This suggests that once the students were in the same grade, there was no difference (an ES -.17 being classified as not significant).

Oldham (1982) retrospectively matched current 11th graders who were retained in primary grades on IQ, gender, and entry age with a group that had been promoted and compared their achievement in reading and math at the 3rd, 6th and 10th grades. This is a same grade comparison. He found that the retained students were significantly higher in mathematics achievement and while higher in reading scores, were not significantly different. The effect sizes were .24, .11 and .11 for reading and .37, .37 and .40 for math at grades 3, 6 and 10 respectively. The average effect size was .28.

Vollrath (1983) compared a sample of those retained in K-3 with a matched sample of students who were recommended for retention but who were promoted instead. He matched on IQ and a cognitive abilities test. He compared their achievement at grades 3 and 6 (same grade comparison) and found that the retained group was significantly higher than the promoted group. On the composite score of the ITBS, the effect sizes were .75 and 1.00 at grades 3 and 6 respectively.

Wright (1979) matched 50 students who had been retained in the first grade with nonretained students on sex, IQ, parent's education and children's achievement scores in the first grade. He carried out a same grade comparison when the children reached the third grade using the CTBS reading, language and math subtests. The average effect size for these same grade comparisons was .29. It is not possible to identify the treatment from the available description.

Collectively, these studies can be classified by educational treatment (remedial vs recycle), by whether the same year or same grade was being compared, and by the time of the comparison (immediately after retention or long term). In Table 2, we indicate the effect sizes for the relevant cell and which study contributed to the cell. It is hazardous to overgeneralize from this table as the number of studies is small. But, this classification suggests the following trends.

INSERT TABLE 2

If we had ignored the type of program and compared students in the same year at the end of the year of retention, the average effect size would have been about zero, leading us to conclude that retention made no difference. However, this conclusion masks the fact that the remedial program showed an average effect size of +.85 while the two studies using recycling had an average effect size of -.86, so that the average appears to be 0.

Certainly the importance of this research lies not in its total effects, but in the finding that different programs have different effects. Other important parts to this story are that the choice of year or grade comparisons will also influence the effect size, as same grade comparisons, quite expectedly, are larger. Finally, because any positive effects of retention are likely to diminish over time, it is important to monitor effects over time.

Discussion

What can be concluded about grade retention from this reexamination? Is the evidence so firm that we can safely say that no child will ever under any circumstances benefit from retention or even that most children will not benefit from retention?

1. Studies which compare students when they are in school for equal time (unequal grade) favor promoted students.

2. Studies which compare students when they are in the same grade (unequal time) favor retained students or show no difference.

3. Studies which present longitudinal comparisons show that any positive effect of retention fades out over a two to three year period.

4. Neither social promotion nor retention per se are effective at solving the problem of providing appropriate instruction for low performing students. The research has been phrased in such a way that a yes or no answer is called for. In fact, the main conclusion should be that both policies are failures. In most cases, doing better than the comparison group still meant a low level of performance relative to the school population at large. Retaining may not help, but simply promoting isn't a solution either.

5. Promotion or retention with additional instruction is more effective than either policy alone.

6. The salient issue is not what policy to adopt (retention vs. social promotion) but how to provide appropriate instruction given student diversity. But this has always been the central question for school organization. Rather than continuing to argue over inappropriate questions, future research should devote attention to locating, developing, and evaluating effective organizational responses to differences in student abilities and competencies. The best that can be said about grade repetition *vs* social promotion as polices is that neither is a victor and that the children are the losers if we continue to debate the issue as it has been formulated thus far.

References

Baenen, N. R. (1988). *Perspectives after five years -- has grade retention passed or failed?* Austin, TX: Austin Independent School District.

Baenen, N. and P. Hopkins. (1988). *New directions: Alternatives to retention, 1987 - 1988*. Austin, TX: Austin Independent School District, Office of Research and Evaluation.

Banerji, R. (1990, April). *A longitudinal study of a developmental kindergarten program.* Paper presented at the annual meeting of the American Educational Research Association, Boston, MA.

Bell, M. (1972). *A study of the readiness room program in a small school district in suburban Detroit, Michigan.* Unpublished doctoral dissertation.

Carlson, R. (1988). *At-risk in kindergarten: A study of selection, retention, transition and promotion.* Unpublished paper, Willamette University.

Coffield. W. H. (1954). *A longitudinal study of the effects of non-promotion on educational achievement in the elementary school.* Doctoral dissertation, State University of Iowa.

Cook, Walter. (1941). *Grouping and promotion in the elementary school.* Minneapolis, Minn: University of Minnesota Press.

Dobbs, V. and D. Neville. (1967, July/August). *The Journal of Educational Research*, 474, 472 - 475.

Farley, E. S. (1936, October). "Regarding Repeaters Sad Effect of Failures Upon the Child." *The Nation's Schools*, 37- 39.

Grissom, J.B. and L.A. Shepard. (1989). "Repeating and Dropping Out of School" in L.A. Shepard *Flunking grades; Research and policies in grade retention.* Falmer Press, New York.

Holmes, C.T. (April, 1986). *A synthesis of recent research on nonpromotion: A five year follow-up.* Paper presented at the annual Meeting of the American Educational Research Association, San Francisco, CA.

Holmes, C. T. and K. M. Matthews. (1984, Summer). "The Effects of Nonpromotion on Elementary and Junior High School Pupils: A Meta-Analysis." *Review of Educational Research*, 225 - 235.

Jackson, G. B. (1975, Fall). "The Research Evidence on the Effects of Grade Retention." *Review of Educational Research*, 613 - 635.

Klene, V. and E. Branson. (1929). (unknown title) *Elementary School Journal* 41:564-566, 1929. Cited in Gregg Jackson. "The Research Evidence on the Effects of Grade Retention." *Review of Educational Research*, 45:613-635, Fall, 1975.

Leinhardt, G. (1980). "Transition Rooms: Promoting Maturation or Reducing Education?" *Journal of Educational Psychology* 72, No. 1, 55 - 61.

Matthews, H. W. (1977). *The effect of transition education, a year of readiness, and beginning reading instruction between kindergarten and first grade.* Unpublished doctoral dissertation, St. Louis University.

Oldham, B.R. (1982). *The longitudinal effects of pupil retention practices in the first three grades.* Unpublished doctoral dissertation.

Peterson, S. E., DeGracie, J. S., and Ayabe, C. R. (1987). "A longitudinal study of the effects of retention/promotion on academic achievement." *American Educational Research Journal, 24*, 1, pp. 107-18.

Raygor, B. (1972). *A five year follow-up study comparing the school achievement variables and school adjustment of children retained in kindergarten and children placed in transition class.* Unpublished doctoral dissertation, University of Minnesota.

Schuyler, N. B. and B. O. Turner. (1986-87). *Retention or promotion: Have policies passed or failed?* Austin, TX: Austin Independent School District.

Shepard, L. A. and M. L. Smith. (1985). *Boulder Valley kindergarten study: retention practices and retention effects.* Boulder, CO: Boulder Valley Public Schools.

- - - (1987, October). "Effects of Kindergarten Retention at the End of First Grade." *Psychology in the Schools*, 346 - 357.

- - - (1988, November). "Escalating Academic Demand in Kindergarten: Counterproductive Policies." *The Elementary School Journal*, 136 - 145.

Shepard, L.A. and M.L. Smith. (1989). eds. *Flunking grades: Research and policies on grade retention*, The Falmer Press, New York.

Smith, M. L. and L. A. Shepard. (1987, October). "What Doesn't Work: Explaining Policies of Retention in the Early Grades." *Phi Delta Kappan*, 129 - 134.

Turley, C. C. (1979) *A study of kindergarten children for whom kindergarten retention was recommended.* Unpublished doctoral dissertation, University of San Francisco.

Vollrath, F. K. (1983). *A comparative study of achievement and classroom behaviors of retained and non-retained kindergarten, third and sixth grade students.* Doctoral dissertation, University of Kansas.

Wright, J. B. (1979). *The measured academic achievement of two groups of first grade students matched along five variables when one group has been retained.* Unpublished doctoral dissertation, Temple University.

Zinski, J. P. (1983). *A study of the effects of a pre-first grade transitional class as compared with first grade retention on reading achievement.* Florida Educational Research and Development Council. Eric ED, 248 459.

Table 1
Retention/Promotion Studies

Author	Sample	Grade of Retn	Match Factors	Effect Measure	Who compared Year/Grade		When compared Time T1	T2	T3	T4	T5	Type of Program	ES
Peterson, deGracie & Ayabe (1987)	1-3 grade retained, promoted	1-3	sex age CAT	CAT			+1	+2	+3			Remedial	+.75 (+)
				RDG	Year	G1	1.06	.24	.30				
				LANG		G1	.41	.29	.19	(+.42)			
				MATH		G1	.83	.35	.10				
				RDG	Year	G2	1.09	.61	.47				
				LANG		G2	.79	.26	.54	(+.62)		=	+.53
				MATH		G2	1.15	.30	.38				
				RDG	Year	G3	.57	.47	.56				
				LANG		G3	.92	1.03	.19	(+.57)			
				MATH		G3	.87	.71	-.15				
				Average			.85	.47	.29				
				RDG	Grade	G1	1.56	.39	.12				
				LANG		G1	1.18	.53	.25	(+.70)		--	
				MATH		G1	2.11	.11	.07				
				RDG	Grade	G2	1.19	1.01	.59				
				LANG		G2	.96	.76	.66	(+.87)		=	+.96
				MATH		G2	1.37	.69	.63				
				RDG	Grade	G3	2.56	.55	.76				
				LANG		G3	2.82	.96	.88	(+1.33)			
				MATH		G3	1.78	.94	.72				
				Average			1.72	.66	.52				
Baenen, N. (1988)	243 retained in 1st grade and matched promoted low achievers	1	ITBS sp ed free lunch age sex	RDG	Grade	G1	.83	.29	.12	-.15	-.26	Remedial	+.15 (NS)
				MATH		G1	.52	.30	.05	-.10	-.11		
							.68	.30	.09	-.13	-.19		

Author	Sample	Grade of Retn	Match Factors	Effect Measure	Who compared	Year/Grade	When compared Time T1	T2	T3	T4	T5	Type of Program	ES
Dobbs & Neville (1967)	60 matched retained/ promoted (30 each)	1	race sex, SES age, MA reading ach	Matro Rdg Math	Year	G1	- .83 -1.07	-.76 -.33		(-.75)		Recycle	(-)
				Rdg Math	Grade	G1	2.67 1.19	1.00 .58		(+1.36)		Recycle	(+)
Coffield (1954)	current 7th graders n=147 students failed & matched pair	3-7	Subtest ITBS in year prior retention	ITBS ITBS	Year Grade		-.77 -.17					Recycling	- (0)
Oldham (1982)	current 11th graders who were retained in primary & retrospectively matched pair n=49	primary	IQ gender entry age	ITBS Rdg Math	Grade		3 .24 .37	6 .11 .37	10 .11 .40	NS p<.05			+.27 +.27
Vollrath (1983)	sample of those retained in K-3 compared with recommended not retained	K-3	IQ CAT	ITBS composite rdg lang math vocab	Grade		3 .75 .60 .41 .71 .64	6 1.00 .44 .67 .68 .34					+.70
Wright (1979)	middle class suburban children; 50 retained in first matched with 45 not retained	G1	Sex IQ Parent's ed Ach scores in 1st	CTBS	Grade	Reading Lang Math	3 .53 .42 .39						+.29

Table 2

Effect Sizes for Retention/Promotion Studies Classified by Comparison Basis, Educational Treatment, and Immediacy of Measurement

	Same year comparisons				Same grade/comparisons			
	Time not specified	+ One Year	+ Two Years	+ Three Years	Time not specified	+ One Year	+ Two Years	+ Three Years
Remedial		.85 a	.47 a	.29 a		1.72 a	.66 a	.52 a
						.68 b		.09 b
Recycle	-.77 d	-.95 c	-.55 b		-.17 d	1.93 c	.79 b	.45 g
					.28 e			
					.63 f			

a Peterson, deGracie, Ayabe (1987)
b Baenen, N (1988)
c Dobbs and Neville (1967)
d Coffield (1954)
e Olaham (1982)
f Voelrath (1983)
g Wright (1979)